

「コンピュータ画像処理による母子相互作用の研究」

石 井 威 望 (東京大学工学部産業機械工学科)

序 論

話し言葉に対応して、動作、表情を同調させる現象をエントレインメントと呼ぶ。その最も原始的かつ根本的な情報交換形態として母子相互作用があげられる。1974年、ボストン大学における研究報告(文献1)によれば、平常な新生児数十名に対して、成人の会話における新生児の体動分析の結果、成人の会話の音節の変化点と新生児の体動の変化点の間に相互同期(Interactional Synchrony)が存在することが示されている。また、米国人の新生児に対して中国語で呼びかけを行なった場合にも同様な結果が得られたことが示され、言語による差異なく、新生児が成人の会話の音節に同期して何らかの意味のある動作をすることが示唆されている。

本研究では、従来定性的色彩の強いこの分野に、新たに画像解析の手法を導入することにより、定量的かつ厳密な科学的評価を試みた。新生児が有意の言語にのみ同期して体動を起こすということが確かであるならば、人間の遺伝情報として社会システムの言語的構造の基礎がすでに新生児期から存在することが推定される。一方、機械文明の重要課題としてマン・マシン・インターフェイスの研究を行なう際にも、本研究はマン・マシン・インターフェイスの原点を与えるのではないかと考えられる。

実験方法

図1に本研究で開発した測定系統図を示す。研究対象の新生児の状態は裸体で、クベース内(温度30℃)に仰臥位で寝かされている。室内の温度が25℃以上ではクベースの外に出して実験を行なう。母親が呼びかけを行なう際には、新生児に触れない。新生児に与える音声刺激は、

1. 母親による自由な呼びかけ(Free Talk)
2. 母親による決まった言葉の呼びかけ、例えば、ハイママヨ、ヨシヨシ、イイコネ(Pattern Talk)

3. 母親以外の人間による自由な呼びかけ及び決まった呼びかけ(Stranger Talk)
4. 計算機合成音
 - a. 周波数、振幅を乱数によって発生させた計算機合成音(Random Noise)
 - b. 人間の呼びかけを模した合成音
 - c. bの合成音を0.1秒のユニットで分割し、順序をランダムに変更して合成した音声のように多種にわたっている。実験の対象としては生後2~4日の正常な新生児を選び、その新生児の1ヶ月後についても追跡調査している。呼びかけ(音声刺激)と新生児の姿はビデオテープに収録する。時間情報はビデオタイマをスーパーインポーズ手法で1/100秒までの時刻をデジタル表示の形式でビデオカセットテープレコーダに画面入力する。さらに今後の開発として、母親の生理的状态を把握するために、計測用マイクロコンピュータを用いて、母親の心電図(R-R間隔)を同時計測する計画であり、目下準備中である。

解析手法及び結果

1. 人間の肉眼による解析

VTRに収録された新生児の体動を人間の肉眼で身体各部(頭、手、足)を個別に動作分析する。動作分析は文献(1)の手法にならない、動作の変化時点を画面上、タイマの値により求める。図2-1に動作分析表を、図2-2に動作分析チャートの一例を示す。

音声については、ビジュグラフの出力波形から音節に対応した時刻を求める。図3にビジュグラフの出力波形の一部を示す。画像との同期は画像上の豆球のON-OFF時のタイミングを記録することにより実現する。

音節と新生児の体動との相関を究明するため、テンプレート・マッチング法を用いて、音声パターンと動作パターンと時系列的にどの程度当てはまるかを大型計算機を使って統計的処理を行なった。図4-1に実データによるテンプレート・マ

マッチングの結果図を示す。図4-2に乱数を導入した動作パターンとのテンプレート・マッチングの結果図を示す。図4-1, 図4-2より実データのグラフの起状の方が乱数データのグラフの起状より大きい。従って、母親の呼びかけと新生児の体動とは何らかの相関があると言える。

また、新生児の体動をMarkov過程であると仮定し、新生児の身体各部の状態推移行列を音節の切れ目のある場合とない場合とについて求めた。その結果の一例を図5に示す。これより、対角要素の総和は音節の切れ目のある部分の方が音節の切れ目のない部分の方より少ない。従って、音節の切れ目のある期間に新生児の体動が頻繁に変化していると言える。

2. 計算機画像処理による自動解析

前述の解析法においては、新生児の体動の変化点や音節の切れ目の検出は人間の肉眼により判断している。従って客観性に欠け、しかも解析に要する労力も過大である。

ここでは、自動的かつ客観的に定量化するために以下の手法を導入した。体動も音声もスカラー量とみなし、動作の変化点のかわりに体動の大きさを、音節の切れ目のかわりに音声の大きさを考える。具体的には、特定の足または手を解析対象として設定し、画面中の設定領域について各フレームごとにデジタル画像(16×16, 明暗度1ビット)に変換し、計算機の記憶領域に格納する。変換の際の閾値は手または足と背景との区別が明確になるように、試行錯誤により決定する。格納されたデータ中で、ある時刻 t における画像と $t + \frac{1}{60}$ 秒における画像との間で対応する画素について、その明暗度の変化のあったものの総数を $x(t)$ とし、体動の大きさを表わすパラメータとする。また音声についてはビデオテープ再生時にその電圧変換値を10 KHzの周波数でサンプリングA/D変換し、 $\frac{1}{60}$ 秒間内での最大値と最小値の差をその時刻 t における音声の大きさ $y(t)$ とする。図6に生後4日の新生児の右手部の動き $x(t)$ と母親の決まった呼びかけ $y(t)$ を示す。この $x(t)$, $y(t)$ から相互相関数 $c(\tau)$ を

$$c(\tau) = \frac{1}{T} \int_0^T (x(t) - \mu_x) (y(t+\tau) - \mu_y) dt$$

により種々の τ について求め、さらに x , y の標

準偏差で除して無次元化する。ここに T は記録長を μ_x , μ_y はそれぞれ x , y の平均値である。図6の例について $c(\tau)$ を求めた結果図(以下、相互相関図と呼ぶ)を図7に示す。この例のように決まった呼びかけでは、音声に対し約1秒遅れて体動が大きく現れている。図8に同一の新生児について自由な語りかけに対する体動との相互相関図を示す。ここでは音声が体動に対し、約1秒早い時点と遅れた時点との2点に大きな山が存在している。音声が先の場合については、決まった語りかけに対して得られた結果と同様、母親の呼びかけに対し体動が同期していることを示している。また音声が後の場合については、母親が新生児の体動に誘起されて呼びかけを行なう結果生じたものと考えられる。

図9に生後4日の新生児についてノイズに対する体動との相互相関図を示す。本図より明らかなように前述したような特徴的な波形のパターンは得られず、母親の呼びかけの場合とは根本的に異なることがわかる。

結論及び展望

本研究により、新生児が母親の呼びかけを他人の呼びかけ、計算機合成音と識別し、それに固有の反応を示していることが示唆された。しかしながら、現在のところ被験者数は十数例と少く、今後2~3年にわたり多数の新生児について、生後2~4日、1ヶ月、2ヶ月と時系列的に追跡調査する予定である。また、心電図計測によるR-R間隔の解析をはじめ、並列的な観測データを利用することによって、実験時の母子の内部状態をより適切に把握することが望まれる。

文 献

- WILLIAM S. CONDON, LOUIS W. SANDER : Neonate Movement is Synchronized with Adult Speech ; Interactional Participation and Language Acquisition, SCIENCE, VOL. 183, 1974.

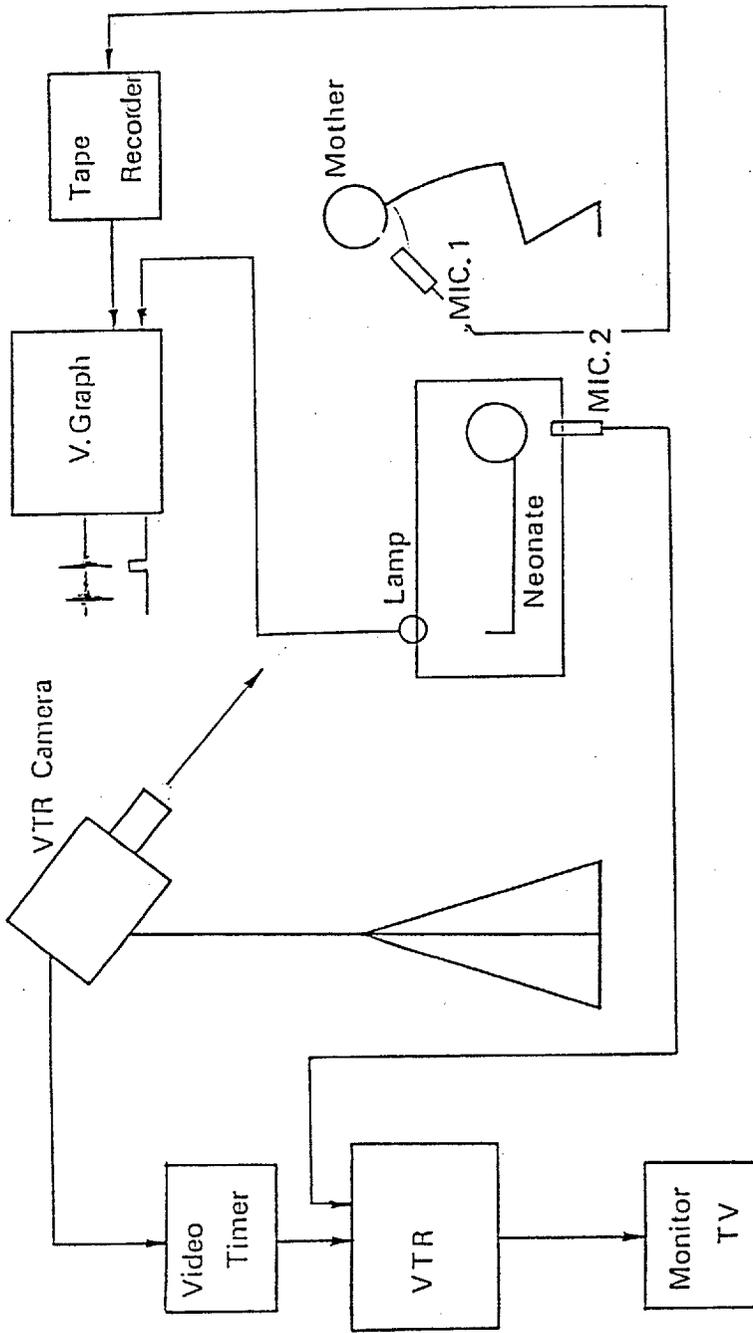
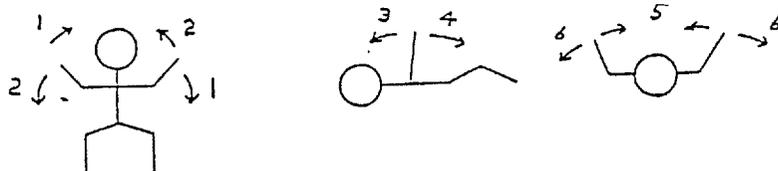


图 1. 测定系统图

新生児の基本動作(上肢)



1. Rotate Clockwise (R.C.)
2. Rotate Counterclockwise (R.Co.)
3. Rotate toward Head (R.H.)
4. Rotate toward Leg (R.L.)
5. Rotate Inward (R.I.)
6. Rotate Outward (R.O.)
7. Rotate Clockwise with Elbow Up (R.C.U.)
8. Rotate Counterclockwise with Elbow Up (R.Co.U.)
9. Elbow Up (E.U.)
10. Elbow Down (E.D.)
11. Stop

図 2-1 動作分析表

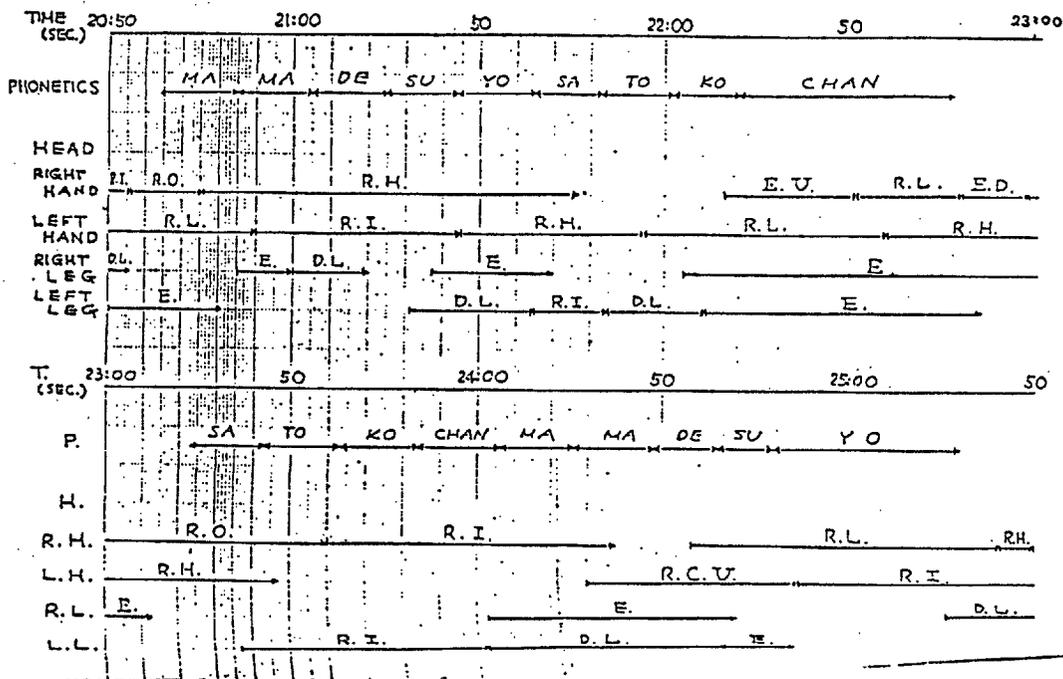
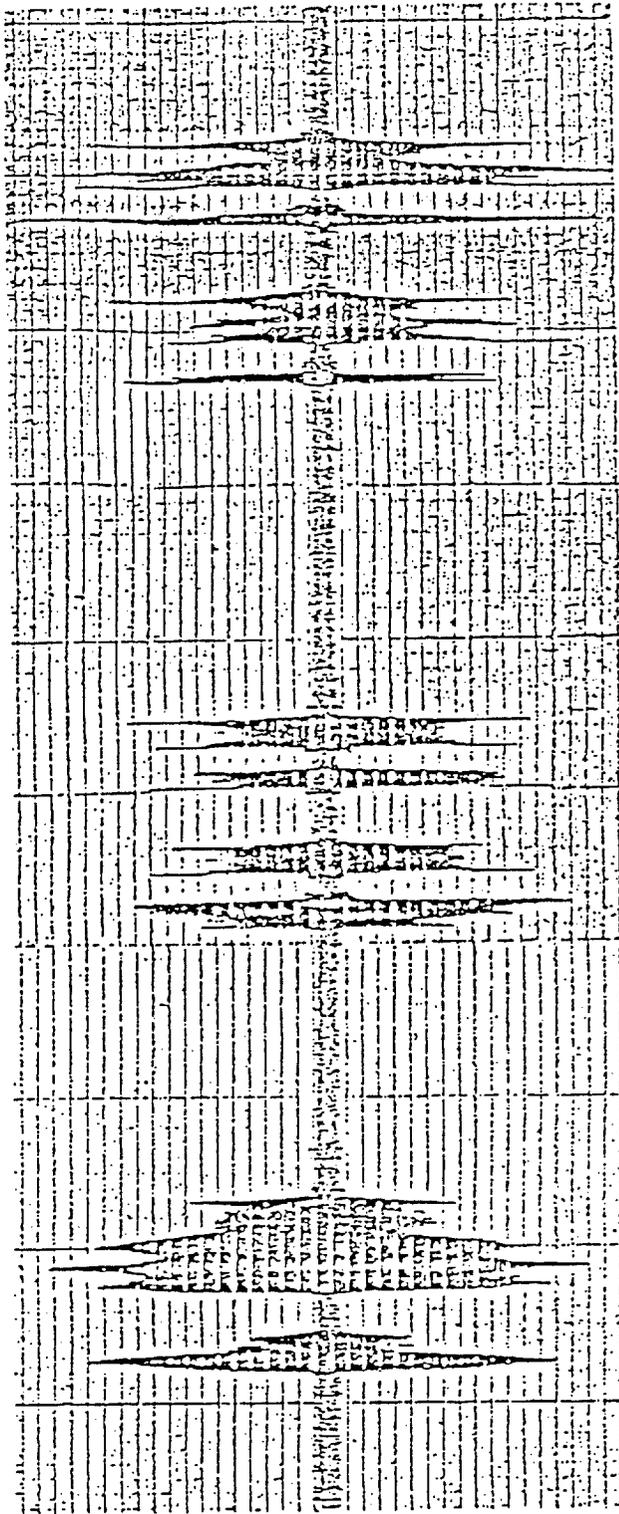


図 2-2 動作分析チャート



おぎたの

おぎたの

ぼく ぼく

はい ママよ

図 3. ビジグラフ出力音声波形 (例)

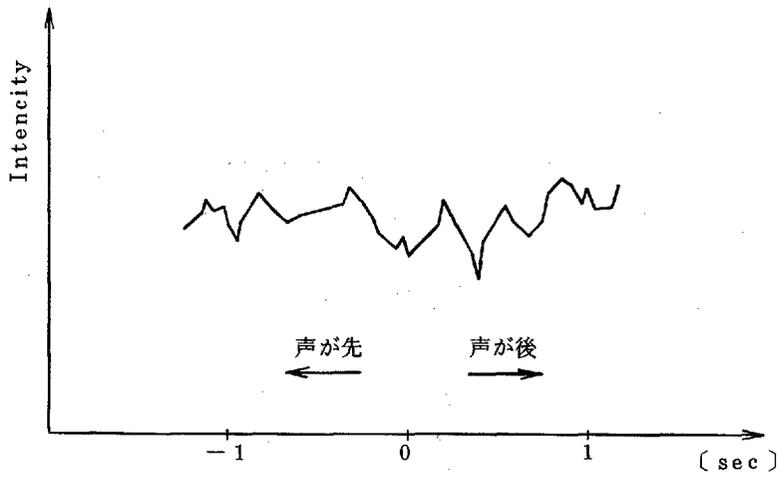


図 4-1 実データ

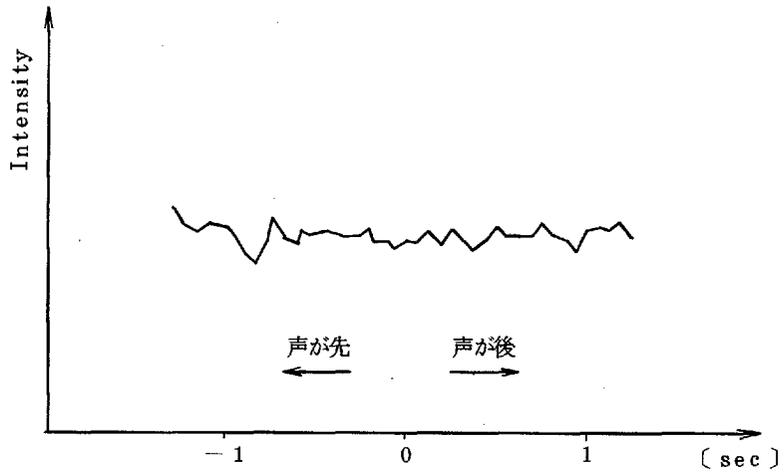


図 4-2 乱数データ

図 4. テンプレートマッチングの結果図

右足

	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
1	94	2	0	0	0	2	3	0	1 <td>71</td> <td>29</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td>	71	29	0	0	0	0	0	0
2	1	98	0	0	0	0	1	0	2 <td>8</td> <td>69</td> <td>0</td> <td>0</td> <td>0</td> <td>8</td> <td>15</td> <td>0</td>	8	69	0	0	0	8	15	0
3	0	0	93	0	0	0	7	0	3 <td>0</td> <td>0</td> <td>0</td> <td>100</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td>	0	0	0	100	0	0	0	0
4	0	0	0	90	0	5	5	0	4 <td>0</td> <td>0</td> <td>0</td> <td>100</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td>	0	0	0	100	0	0	0	0
5	0	3	0	0	94	0	3	0	5 <td>0</td> <td>100</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td>	0	100	0	0	0	0	0	0
6	0	0	0	0	0	81	14	5	6 <td>8</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>84</td> <td>0</td> <td>8</td>	8	0	0	0	0	84	0	8
7	1	1	0	0	0	0	98	0	7 <td>0</td> <td>4</td> <td>0</td> <td>4</td> <td>0</td> <td>0</td> <td>92</td> <td>0</td>	0	4	0	4	0	0	92	0
8	0	0	0	0	0	4	4	92	8 <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>100</td>	0	0	0	0	0	0	0	100

左：音節の切れ目のない部分
 右：音節の切れ目のある部分

図5. 推移確率行列

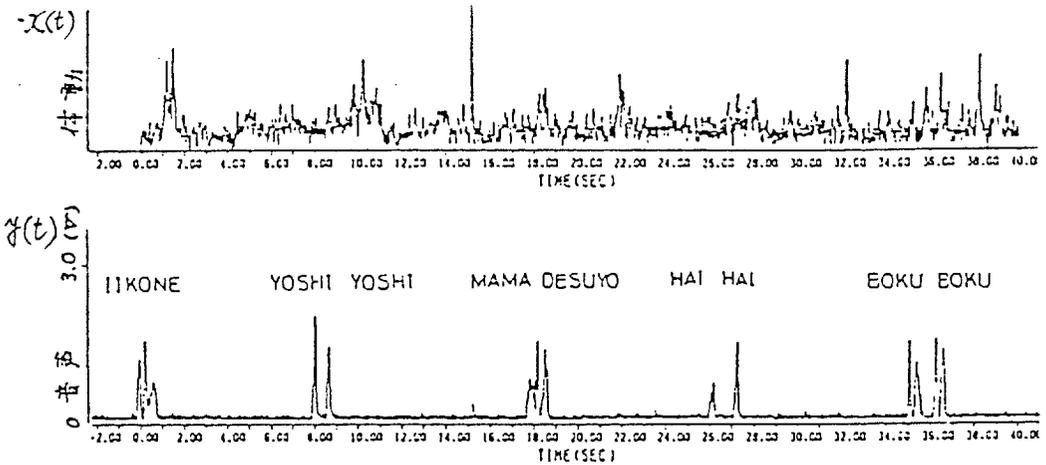


図6. 生後4日男子右手部体動と、母親の音声

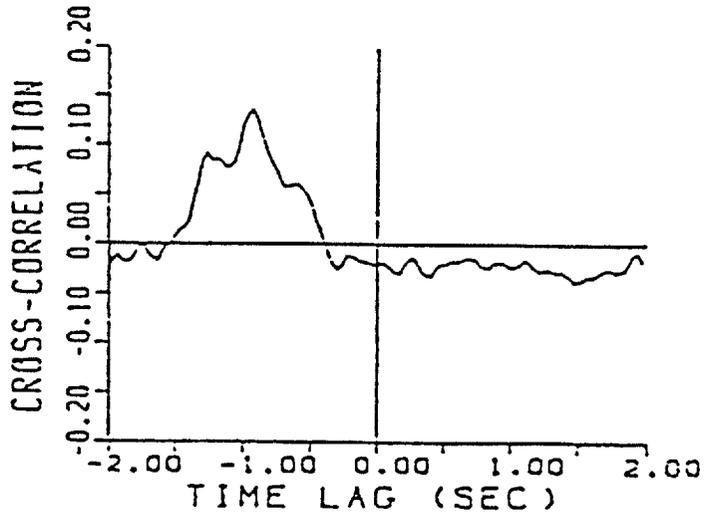
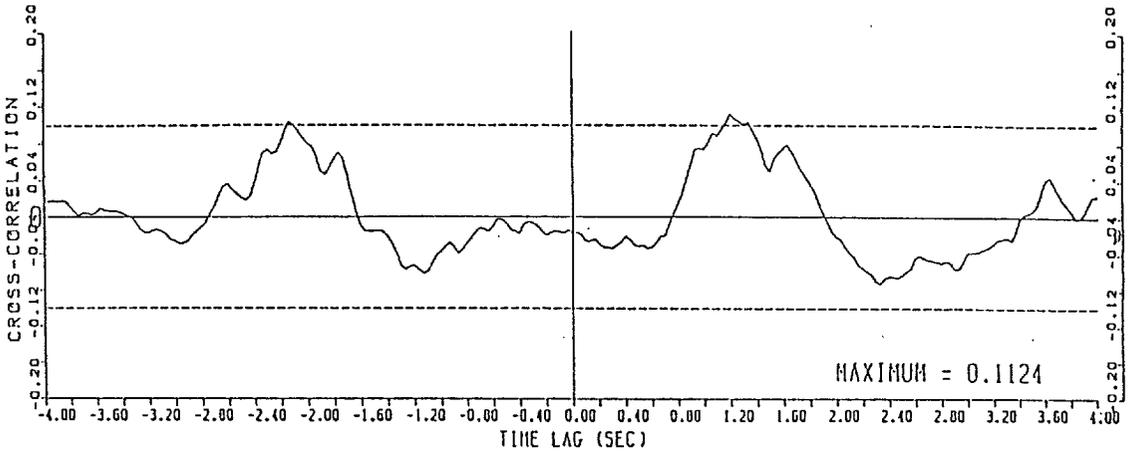
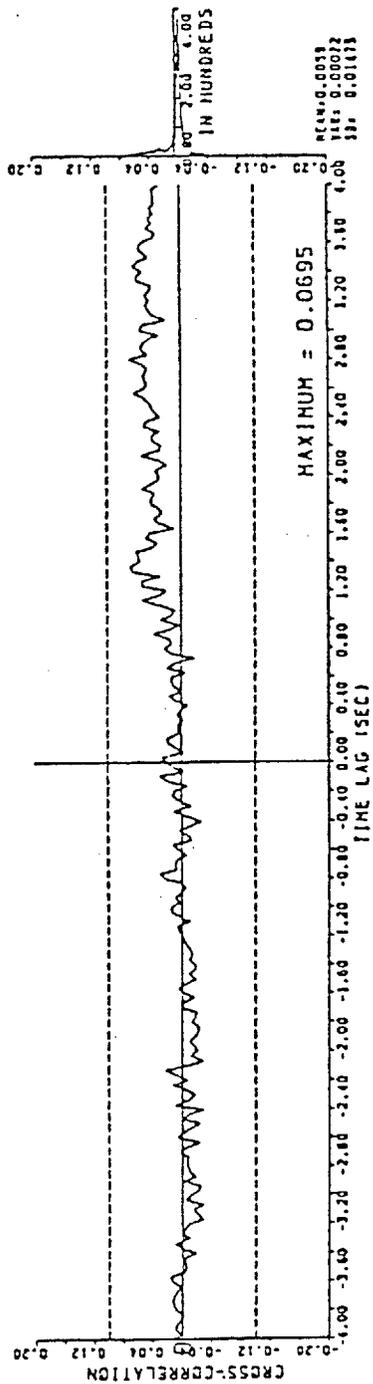


図7. PATTARN TALK 相互相関図



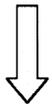
4DAY-OLD MALE 9/30 FREE TALK-HAND

図8. FREE TALK 相互相関図



1 MONTH FEMAF 6/26 NOISE HAND RIGHTLEG4 LEG

图 9. RANDOM NOISE 相互相关图



検索用テキスト OCR(光学的文字認識)ソフト使用

論文の一部ですが、認識率の関係で誤字が含まれる場合があります



序論

話し言葉に対応して、動作、表情を同調させる現象をエントレインメントと呼ぶ。その最も原始的かつ根本的な情報交換形態として母子相互作用があげられる。1974年、ボストン大学における研究報告(文献 1)によれば、平常な新生児数十名に対して、成人の会話における新生児の体動分析の結果、成人の会話の音節の変化点と新生児の体動の変化点の間に相互同期(Interactional Synchrony)が存在することが示されている。また、米国人の新生児に対して中国語で呼びかけを行なった場合にも同様な結果が得られたことが示され、言語による差異なく、新生児が成人の会話の音節に同期して何らかの意味のある動作をすることが示唆されている。

本研究では、従来定性的色彩の強いこの分野に、新たに画像解析の手法を導入することにより、定量的かつ厳密な科学的評価を試みた。新生児が有意の言語にのみ同期して体動を起こすということが確かであるならば、人間の遺伝情報として社会システムの言語的構造の基礎がすでに新生児期から存在することが推定される。一方、機械文明の重要課題としてマン・マシン・インターフェイスの研究を行なう際にも、本研究はマン・マシン・インターフェイスの原点を与えるのではないかと考えられる。